International productivity comparisons and natural resources: resource rents and missing endowments

Freeman, Daan; Inklaar, Robert

GGDC RESEARCH MEMORANDUM *178*

International productivity comparisons and natural resources: resource rents and missing endowments

Daan Freeman and Robert Inklaar

June 2018

university of groningen

groningen growth and development centre

International productivity comparisons and natural resources: resource rents and missing endowments

Daan Freeman[1] and Robert Inklaar

Groningen Growth and Development Centre, University of Groningen

June 2018

Abstract

Standard theory for cross-country productivity comparisons assumes all countries use the same set of factor inputs in production. This assumption is violated when extending the set of (potential) factor inputs to include natural resources, such as oil, gas and gold, because countries typically lack certain endowments. Yet the extension of factor inputs beyond produced capital and labour is important for arriving at unbiased relative productivity estimates. In this paper we propose a solution to the missing endowment problem by comparing productivity only for the overlapping set of inputs *and activities*. We show that this has a substantial impact on estimated relative productivity levels for those countries heavily reliant on natural resources for generating their income.

JEL codes: E22; O13; O57; O47; Q32

Keywords: Productivity measurement; natural resources; development accounting

[1] Corresponding author; (d.freeman@rug.nl), Groningen Growth and Development Centre, Faculty of Economics and Business, University of Groningen, Nettelbosje 2, 9747 AE Groningen, The Netherlands.
Acknowledgments: We would like to thank participants at the SOM PhD conference (2017) and Society for Economic Measurement conference (2017) for helpful comments.

Introduction

Development accounting is a popular tool that is used to establish how much of the differences in income levels across countries can be accounted for by differences in observed factor inputs – such as buildings, machinery and (skilled) workers – and how much to differences in productivity, the residual.[2] This, in turn, can then inform further research to explain why investment in capital may be low or why productivity lags.[3] But omission or mismeasurement of factor inputs will lead to biased measures of productivity. This has motivated researchers to expand and improve the measurement of inputs, by including additional types of intangible capital (Chen, 2017), accounting for differences in management practices (Bloom, et al., 2016) and improving estimates of human capital over the life cycle (Lagakos, et al., 2017, Inklaar and Papakonstantinou, 2018). Omitted so far in these efforts is the role of natural resources,[4] such as oil, gas, iron and gold, even though natural resources are an important source of income and wealth in many lower-income countries, as well as some (very) high-income countries (Lange, Wodon and Carey, 2018).

The contribution of this paper is to propose and implement a method for incorporating natural resources as a factor of production in cross-country comparisons of productivity. We build on the work of Brandt et al. (2017), and Diewert and Fox (2016), who show how natural resources can be incorporated in a 'sources of growth' framework. Many of the measurement considerations of their work, such as measures of resource rents, carry through to a cross-country context. However, the extension to a cross-country setting

---

[2] See Caselli (2005) and Hsieh and Klenow (2010) for overviews of this literature.
[3] See e.g. Acemoglu et al. (2016), who show that democratization increases income levels by improving investment, not by improving TFP.
[4] In National Accounts terminology, these are typically referred to as 'subsoil assets'.

faces a notable challenge in that most countries lack an endowment of at least one type of natural resource. Such missing endowments mean that relative productivity is not defined in the typical productivity comparison framework, such as that of Diewert and Morrison (1986) and Inklaar and Diewert (2016).

We propose a solution, namely to compare 'like with like'. If country A has no (economically viable) endowments of natural resource asset X, this means it also lacks a mining industry devoted to the extraction of asset X. If country B does have an endowment of asset X, we cannot compare the productivity of economy A to the *entire* economy B as the input of X in country A equals zero. However, we can compare A to a truncated version of economy B, one that excludes the input *and industry* associated with asset X. This exploits the feature of (this class of) natural resources that only a mining industry can utilize a natural resource as an input. After extraction, the ore, liquid or gas is a regular product that can be used in other industries, domestic and foreign.

We illustrate this method for incorporating natural resources in international productivity comparisons for a sample of countries. We select countries with diverse endowments, including some with endowments of all or nearly all resources (Australia, Canada, United States), those with a few resources and a large income from these resources (Saudi Arabia, Qatar), and some with few resources and small associated incomes (Japan, Portugal). This allows us to illustrate, firstly, how estimates of relative productivity are affected by the importance of natural resources as an income source, and secondly, the impact on productivity estimates of the missing endowment problem.

Our results show that incorporating natural resources is non-trivial for the productivity estimates of several countries, particularly those with larger endowments relative to GDP. The results indicate that for several resource rich economies, the inclusion natural resources into the accounting equation means a substantial correction to estimated productivity figures. Out of the countries we consider, the estimated productivities of Qatar, Saudi Arabia, and Australia altered to most. These three countries have the largest shares of natural resource income to GDP. Secondly, we estimate indirect productivity differences using a third country as an intermediate link. We use the indirect comparisons when direct comparisons cannot account for all endowments. Using our truncation approach, we find substantial differences between our direct productivity estimates, and the indirect productivity estimates. These differences are based on the degree to which the structures of the different economies overlap; the differences illustrate the difficulty of accurately comparing countries with very different production structures.

This research relates to several strands of literature. First, compared to standard development accounting approaches, we extend the scope of the capital concept, similar to the contribution of Chen (2017) and, more broadly, related to the literature that has emphasised that the scope and contribution of capital inputs can reveal a more prominent role for capital in accounting for cross-country income differences, see e.g. Caselli & Feyrer (2007) and Mutreja (2014).

This paper is also related to index number literature, specifically, the literature that deals with the 'new goods' problem. New goods are a particular challenge to standard index number theory, because the new (or disappearing) good is not available in one of the periods (or countries) under consideration. Various approaches exist for dealing with the new goods problem; Balk (1999) reviews several methods, and Redding & Weinstein (2018), extending the method developed by Feenstra (1994), provide a recent contribution to this discussion. In our productivity measurement framework, the insights from this literature are not directly applicable since the 'new goods' literature provides solutions for adjusting the *price index.* However, productivity measurement theory prescribes an input *quantity* index rather than a price index and an input quantity of zero is fundamentally problematic in a way that a 'new goods' adjustment is not feasible. One key insight that carries over from the 'new goods' literature is that the price index should be based on products observed in both periods. This corresponds to our approach of comparing 'like with like', so only comparing the productivity for the part of the economies with overlapping endowments of natural resources.

The rest of this paper is structured as follows, first the development accounting methodology is outlined. Subsequently, the problem arising from the incorporation of natural resources is described in more detail, after which our truncation approach is outlined. This is followed by a section introducing the data, which also presents the results to illustrate our approach by applying it to a number of countries with different endowment sets. A final section concludes.

Methodology

This section draws on Inklaar & Diewert (2016) – along with Diewert & Morrison (1986) – and reviews the main methodology for international productivity comparisons. Subsequently, the shortcoming of the method in the case of missing endowments is discussed, and our contribution is presented.

As in Inklaar & Diewert (2016), consider a set of $i = 1, \dots, I$ production units[5], each endowed with a strictly positive[6] vector $x \equiv [x_1, \dots, x_N] \gg 0_N$ containing N input factors, and producing a vector $y \equiv [y_1, \dots, y_M]$ containing M *net* outputs[7]. Additionally, assume a

[5] These production units can be defined at different levels; e.g. firms, industries, broad sectors.
[6] For now, changing this assumption motivates the latter parts of this section.
[7] Where positive values indicate goods produced; negative values indicate the good is an intermediary input.

strictly positive net output price vector $p \equiv [p_1, \dots, p_M] \gg 0_M$. Now, the value added function $g^i(p, x)$ for each production unit i is defined as:

$$g^i(p,x) \equiv \max_y \left\{ \sum\nolimits_{m=1}^{M} p_m y_m : (y, x) \in S^i \right\}; i = 1, \dots, I \tag{1}$$

Equation (1) shows that g^i is the maximum value of output in production unit i given S^i; the feasible set of inputs and net outputs, subject to constant returns to scale. Diewert & Morrison (1986) use the translog functional form to specify the value added function $g^i(p, x)$:

$$\ln g^i(p,x) \equiv \alpha_0^i + \sum_{m=1}^{M} \alpha_m^i \ln p_m + \frac{1}{2}\sum_{m=1}^{M}\sum_{l=1}^{M} \alpha_{ml} \ln p_m \ln p_l + \sum_{n=1}^{N} \beta_n^i \ln x_n + \frac{1}{2}\sum_{n=1}^{N}\sum_{o=1}^{N} \beta_{no} \ln x_n \ln x_o + \sum_{n=1}^{N}\sum_{m=1}^{M} \gamma_{nm} \ln x_n \ln p_m \tag{2}$$

In the presence of several other assumptions[8], the translog function is well-suited for productivity comparisons between two production units. Given this specification of the value added function, Diewert & Morrison (1986) derive a consistent bilateral productivity index between two countries.[9] This index requires several building blocks that will be now be discussed following Inklaar & Diewert (2016).

First, define the value of each net output as v_{km} for each net output $m = 1, \dots, M$ and country $k = 1, \dots, K$. Likewise the value of each input factor is V_{kn}, for each input factor $n = 1, \dots N$ and country $k = 1, \dots, K$. Having defined these values, the share of each net

[8] Note that the parameters α_0^i, α_m^i, and β_n^i with country-superscripts are allowed to vary across countries, but α_{ml}, β_{no}, and γ_{nm} are country invariant. Additionally note that $\alpha_{ml} = \alpha_{lm}$ and $\beta_{no} = \beta_{on}$. Further parameter restrictions apply ensuring the linear homogeneity of g^i (see Diewert (1980); Diewert & Morrison (1986).

[9] This bilateral case is extended to the case multilateral by Inklaar & Diewert (2016); however, due to the nature of the current work, construction of a multilateral index is not feasible so this discussion is omitted here.

output (input factor) in the value of total country net outputs (input factors) can be defined as:

$$s_{km} \equiv v_{km}/v_k\,;\; m = 1, \dots, M;\; k = 1, \dots, K \quad (3)$$

$$S_{km} \equiv V_{kn}/V_k\,;\; n = 1, \dots, M;\; k = 1, \dots, K \quad (4)$$

Where $v_k \equiv \sum_{m=1}^{M} v_{km}$ and $V_k \equiv \sum_{n=1}^{N} V_{kn}$ are the total value of net outputs and input factors for each country k.

Additionally, define p_{km} as the purchasing power parity (PPP) price for net output m in country k, and w_{kn} as the PPP price for input factor n in country k. The relevance of using PPP prices is that they provide a consistent measure across different countries for the same net output or input factor. Finally, by combining the information on the values and PPPs, the (implicit) quantities can be estimated. Specifically, $y_{km} \equiv v_{km}/p_{km}$ is the quantity of net output m, and $x_{kn} \equiv V_{kn}/w_{kn}$ is the quantity of input factor n, both in country k.

Having defined the values (v_{km} and V_{kn}), the value shares (s_{km} and S_{kn}), the prices (p_{km} and w_{kn}), and the quantities (y_{km} and x_{kn}) for each net output m and input factor n we can define the Diewert & Morrison (1986) productivity index $\Gamma_{k/j}$, comparing countries k and j for $k, j = 1, \dots, k$.

$$\Gamma_{k/j} \equiv Y_{k/j}/X_{k/j}; k, j = 1, \dots, K \quad (5)$$

This equation is often used in the literature to make bilateral productivity comparisons (see for example Feenstra et al, 2015). The productivity index, $\Gamma_{k/j}$ is made up of two distinct indexes, the output quantity index $Y_{k/j}$ and the input quantity index $X_{k/j}$:

$$Y_{k/j} \equiv [v_k/v_j]/\exp\left[\sum_{m=1}^{M}\frac{1}{2}(s_{jm}+s_{km})\ln(p_{km}/p_{jm})\right]; k,j=1,\dots,K \tag{6}$$

$$X_{k/j} \equiv \exp\left[\sum_{n=1}^{N}\frac{1}{2}(S_{jn}+S_{kn})\ln(x_{kn}/x_{jn})\right]; k,j=1,\dots,K \tag{7}$$

These can each be defined using the variables introduced above, allowing us to specify the productivity function in an applied setting. However, recall that one assumption necessary for this method to work is that the vector of inputs x is strictly positive. This is also obvious from equation (7), as the natural logarithm in the final term is only defined for positive values of the relative input factors, i.e. $x_{kn}/x_{jn} > 0$. This means that when comparing two countries, their non-zero inputs need to be the same set of inputs (a missing input in one country is in effect a zero value of that input the input vector). Therefore, a problem arises when many different inputs, some of which missing in particular countries, are considered. This problem we refer to as the missing endowments problem and finding a way around it is the primary objective of this paper. A more detailed discussion continues in the next section.

The missing endowments problem

The problem arises when bilateral productivity comparisons feature unequal sets of factor inputs. Consider again $i = (1, \dots, I)$ production units, each using a set (or vector) $x_i = [x_{i1}, \dots x_{iN}]$ of input factors. Assume at least a subset of these input factors geographically fixed; specifically, that they cannot cross borders[10]. This means that when

[10] While such an assumption does not hold for many factors that are man-made, like machines, trucks and computers, it is true for many natural resources. For example, a stock of subsoil ore cannot be moved abroad without first being extracted (at which point it becomes a commodity, and is no longer considered an input factor).

comparing the same production unit across two countries; k and j, their sets of inputs do not correspond, i.e. $x_k \neq x_j$ (having dropped production unit subscripts). As such, a complete comparison between the two countries is not possible with the Diewert & Morrison (1986) methodology.

Consider first that if $x_k = x_j$, the comparison can be made readily, as the sets of inputs used in both countries overlap, and can be compared consistently using equation (5). Alternatively, if $x_k \neq x_j$, the endowment sets are no longer the same, and a comparison between countries k and j runs into trouble. Assuming however, that there is at least some overlap in endowments, there is a set of inputs x_{kj} that both countries have in common. The set x_{kj} is therefore a subset of both x_k and x_j, as each country might be endowed with additional inputs that the other country does not have, or more formally, $x_{kj} = (x_k \cap x_j) \neq \emptyset$.

A special case of this situation is when x_{kj} equals either x_k or x_j. In this case, the endowments of one country do not just differ from those of the other, but also completely eclipse them. For example, if $x_k \in x_j$ holds, the set of inputs of country k is overshadowed entirely by those of j. In this case the common set x_{kj}, equals the endowments of k; $x_{kj} = x_k$. Without loss of generality, this situation will be used throughout the subsequent sections.[11]

The bias

If we would ignore the problem of missing endowments, and use equation (5) to evaluate productivity differences between two countries without additional consideration, we

[11] The method presented in this paper relies on some overlap in endowments, which is ensured by the presence of labour and fixed assets as part of factor endowments.

would introduce a bias in the productivity estimates.[12] Consider again countries k and j, featuring otherwise identical endowments, with the exception of input $\chi \in x_1, \dots, x_N$. Country j is endowed with χ, but k is not, or more specifically $\chi_k = 0$, while $\chi_j > 0$. As suggested above, if equation (7) is used, and input χ remains unconsidered, a part of income is left unaccounted for. What happens to the estimates of productivity differences if we ignore this?

Since a part of income is not being accounted for, we violate our assumption of constant returns, which states that $V_j = v_j$, or that the value of total inputs (V_j) equals the value of total outputs (v_j). This is because, if we do not take factor χ_j into account, we artificially lower value of inputs vis-à-vis the value of outputs; i.e. we end up with a situation in which $V_j < v_j$[13]. Unfortunately, the model is not equipped to deal with this situation, which by equation (4) enforces constant returns through a unity sum of input shares; $\sum S_{jn} = 1$. As factor χ_j is left out of consideration, and by assigning the income share of factor χ to other factors, V_j (without factor χ_j) is `stretched' to match v_j. Due to their inflated input shares, the other factors will appear more productive, overestimating the overall productivity estimate of country j. In more general terms, for either country, each non-zero factor that is left out of consideration, introduces an upward bias in its productivity estimate.

[12] Intuitively, a solution to the missing factor endowment problem could be to aggregate factors up to a level at which both countries have at least some value. This procedure would be correct only if the marginal products of the detailed factors are identical. In general, there is no reason to assume that this is the case for different production factors. In our illustration below, we show that assuming equal marginal products across assets is not likely valid.

[13] As in effect, V_j is the value of all inputs, minus $V_{j\chi}$, the value of input χ.

Factor-industry value added

This section presents our approach to addressing the missing endowments problem outlined in the above sections. We aim to estimate '*truncated'* versions of countries' economies, where specific industries and factor inputs are omitted. In terms of countries k and j, with the input $\chi_k = 0$, the aim would be to compare country k to country j if input $\chi_j = 0$. To achieve this, the value that is added by employing input χ_j needs to be subtracted from the total value of outputs, or GDP of country j. This will then satisfy $V_j = v_j$. Instead of stretching the value of all inputs (V_j) through inflated input shares, we will reduce the value of total output (v_j) to bring it in line with the value of inputs without factor χ_j.

Now, suppose input χ is only used by one specific production unit, or industry $z \in I$. Furthermore, assume that total value added generated in country j's industry Z is represented by $v_{Zj} = Y_Z(x_j^z, P_j)$, with $P_j = [p_{j1}, \dots, p_{jM}]$, a vector of prices for M net outputs, and $x_j^z = [x_{j1}^z, \dots, \chi_j^z, \dots, x_{jN_z}^z] \leq x_j$, a set of N_z inputs employed in industry z.[14] Given the value added of industry z_j, we can reduce the value of total output by this value, to arrive at an estimate of the total output of country j, *without* industry z_j, and therefore factor χ_j:

$$\bar{v}_j = v_j - v_{zj} \tag{8}$$

Where the total value of output in country j (v_j) is reduced by the output of industry Z (v_{zj}). Therefore, $\bar{v}_j$ is the value of country j's *truncated* total output (or GDP) and is an

[14] N_z by assumption includes in χ, but also inputs factors general to all industries (capital and labour). Note the lack of country subscript, implying the assumption that industries z use the same factors inputs in all countries.

estimate of country j's GDP had it not been endowed with input χ_j. Equation (8) can readily be extended to subtract the value added of multiple industries $v_{zj}, z \in Z$.

Finally, recall that the vector of inputs employed in any industry z, x_j^z, includes not only factor χ, but also other input factors *general* to all industries, such as labour. Since we reduce GDP of country j by the *total* value generated by industries z, these general input factors also need to be adjusted.[15] To leave the economy-wide inputs of general factors in country j the same, would introduce a bias, as the country-level aggregates of these factor inputs would be overstated by the amounts employed in the removed industries. To avoid this bias, we apply a similar truncation method to the aggregate general factor inputs, as we have to GDP.

Comparing this situation to the case where the missing endowments are simply ignored, using our solution removes the value generated by industries Z, from v_j. This means it is now 'safe' to specify equation (7) without inputs χ, for either country, since the GDP of neither includes any value derived from that input anymore. This alters the interpretation of equation (5): it is no longer the case that we are comparing the productivities of complete economies with each other; but we are comparing country k to a *truncated* country j. The countries are only compared on the basis of inputs and associated outputs of industries that they have in common. Any income derived from, and therefore the productivity of, the other inputs (χ) due to industries Z are left out of consideration.

[15] In practice, these general input factors include any used by all the industries, which in our specification means capital and labour.

Empirical illustration

The missing endowment problem can be illustrated using productivity comparisons that include subsoil natural resources as factor inputs. Countries often lack endowments of one or more of these assets. Additionally, subsoil assets are generally employed in only one industry: the relevant extraction (mining) industry. These facts make subsoil assets ideally suited as inputs χ, to illustrate our approach outlined in the previous section.

Since the focus here is on subsoil assets, geographically large countries generally have the most diverse endowments. Such countries are well suited to serve as a base country (country j in equation (5)); given sufficiently detailed industry data on value added, they can be truncated to match the endowments of any other country. This enables us to compare each country using the same base country, albeit different truncations of this base country.

Data

The description of the data proceeds into three parts, specifically according to equation (6), the output quantity index, secondly, the input quantity index in equation (7), and finally the industry data require for the truncations.

Firstly, the output quantity index requires data for the total value of output, v_k for each country $k = 1, \dots, K$. This is the value of GDP for each country, which is obtained from the PWT 9.0. The second part of equation (6) consist of a price index, which consists of data for the output shares (s_{km}) and output prices (p_{km}), for each net output $m = 1, \dots M$. Such price indexes for the GDP (v_k) of each country are estimated (and used) by Feenstra et al (2015) and are available from the PWT 9.0.

The input index in equation (7) requires data on the inputs shares, S_{kn}, and the quantities of factors, or inputs, x_{kn}, for each input factor $n = 1, \dots N$. The data on factors that we use can be divided into three broad categories; capital, labour (or human capital) and subsoil assets. The capital data is constructed using the Perpetual Inventory Method (PIM), and is available from the latest version of the PWT 9.0 (Feensta et al 2015). The values of w_{kn}, or data on prices (PPPs) to make the stocks comparable across countries, is also available from the PWT 9.0. The PWT 9.0 is also the source for data on human capital, which reflects the size and quality of the labour force. Human capital is based on Mincerian returns, and described in more detail, again in Feenstra et al (2015), and the PWT 9.0 documentation.

Second, the data for subsoil assets are acquired from the World Bank's wealth accounting project (Lange et al., 2018); data on prices, costs and output for different types of subsoil assets are available. To make the data on subsoil assets suitable for equation (7), capital stocks need to be estimated, which is done by estimating the present value of the revenue streams.[16] As stated before, not all countries are endowed with the full range of subsoil assets available in the data. In fact, of the entire World Bank data set, in 2011 only a few countries can be considered endowed with at least some quantity of each of the included factors – see Table 1, below, for the list of assets and coverage across the set of countries we analyse.

[16] The equation used to estimate the stocks is $V_a^t = \sum_{i=t}^{t+T-1} \frac{\bar{R}_a^t}{(1+r)^{i-t}}$. Where V_a^t is the value of an asset a's stock, and R_a^t is the revenue derived from its exploitation, in year t. The bar over R indicates it is a 5-year moving average including t through $t-4$. The real interest rate r is assumed 4%, and T is the lifetime of asset a, both of which are taken from Lange et al. (2018). The constant $\bar{R}^t$ furthermore implies assuming the flow of rents is constant in future periods. These values are expressed in US dollars at current market exchange rates. By applying these values in the current context, this effectively assumes the law-of-one-price holds so that the exchange rate suffices to compare 'real stocks' across countries.

Therefore, the PWT 9.0 data and the World Bank wealth accounting data cover the required data on factor inputs. This leaves only the input shares to complete the specification of equation (7). These shares are approximated by the share of income accruing to each factor included in our specification. It is common in the literature to assume some share of labour income, and setting capital income as the residual (Caselli 2005; Hsieh & Klenow 2007). Specifically, the income share of labour is available for each country from the PWT 9.0, and is measured by the share of income accruing to workers. However, the income shares of capital and subsoil asset factors are more difficult to observe in a similar way. For this reason, the income accruing to subsoil assets is estimated by using World Bank data on prices, costs and production of each of the subsoil assets. These figures can then be used to obtain the income accruing to the subsoil assets, and to derive their shares. This is approach implicitly uses user costs to estimate capital income. In fact, Diewert & Fox (2016) describe two (equivalent) ways in which the user costs of capital can be derived, one for reproducible capital, and another for non-renewable capital. We apply the latter here, as subsoil assets are non-renewable resources.[17] Finally, assuming the income share of capital is a residual allows us to specify equation (5), and to estimate bilateral productivity differences.

[17] The user cost for produced capital is $UC^C = P[r - i + (1+i)\delta]$, Where the real rate of return (r) is assumed to be 4%, price changes (i) are from PWT 9.0, and depreciation rates (δ) from Feenstra et al (2015). The user cost for non-renewable resources capital is $UC^N = R/(\Delta S)$, where R is the net revenue of resource extraction and ΔS the change in the associated stock. Diewert & Fox (2016) furthermore show that, subject to some assumptions, these two user costs are equivalent.

Table 1 – Subsoil asset endowments and income share in 2011

	Gold	Copper	Gas	Oil	Silver	Zinc	Iron ore	Lead	Phosphate	Therm. coal	Metall. coal	Nickel	Brown coal	Bauxite	Tin	Income share
AUS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	9.5%
CAN	1	1	1	1	1	1	1	1	1	1	1	1	1			3.2%
USA	1	1	1	1	1	1	1	1	1	1	1		1	1		1.2%
ZAF	1	1	1	1	1	1	1	1	1	1	1	1				8.2%
SAU	1	1	1	1	1	1			1					1		50.2%
SWE	1	1			1	1	1	1								0.8%
PRT		1			1	1	1								1	0.2%
JPN	1		1	1	1											0.0%
NLD			1	1												0.8%
QAT			1	1												38.2%
URY	1															0.1%

Source: Lange et al. (2018).

Note: Whenever a country has positive production for a subsoil asset, a 1 is entered. The final column shows for each country the net rents derived from their total stock of subsoil assets, as a percentage of GDP.

Table 1 displays the coverage of the various subsoil assets for each country included in this illustrative exercise. A "1" in a cell indicates that a country has endowments of the corresponding asset, listed in the columns. This table gives an idea about the extent to which countries need to be truncated in order to bring their endowments in line with those of other countries. For example, if we want to evaluate the productivity of the Japan (JPN) relative to the United States (USA), the United States needs to be truncated in terms of industries corresponding to the extraction of copper, zinc, iron, lead, all types of coal, phosphate, and bauxite. This leaves gold, oil, gas, and silver, which are the endowments of Japan, and is the set of common factors we can use for the productivity comparison[18].

The final column of Table 1 displays the percentage of each country's GDP that is derived from the use of all the subsoil assets. Some striking differences can be observed;

[18] Of course, in addition to capital and labour factors, which are utilised in all industries.

specifically, that Qatar (QAT) and Saudi-Arabia (SAU) are highly dependent on their subsoil assets, while countries like Portugal (PRT), Japan (JPN), and Uruguay (URY) derive only minor shares of their GDP from subsoil assets. This latter fact is, of course, not a strange thing, as they are endowed with only a limited set of assets, in addition to these being of small quantities.

We select base countries based on their range of available subsoil assets. This leads us to choose Australia, Canada, and the United States as the base countries for the subsequent analyses. Australia is one of the few countries that have complete coverage of the World Bank subsoil assets; additionally, (relatively) detailed industry accounts are available. The United States does not cover all the subsoil assets that the World Bank dataset contains; however, the detail of the industry accounts makes it suitable as another base country. Additionally, we use Canada due to its wide coverage, and relatively complete industry accounts. We will perform the bilateral productivity comparisons using the other countries listed in Table 1, each differently endowed with various subsoil assets, to illustrate our method.

Finally, the third set of data required is the mining industry data, to be able to truncate the base countries to make them comparable to other countries. This requires information on the value added, labour input and capital input associated with each subsoil asset in the three base countries, Australia, Canada and the United States. We combine data from the Socio-Economic Accounts of the World Input Output Database (Timmer et al., 2015) with more detailed estimates of the structure of mining industry in the three countries; Appendix A provides more detailed information.

Results

Using the data presented in the previous section, we can specify the productivity index of equation (5). Table 2 shows the relative productivity estimates for 2011[19], using the USA as the base country. Each cell indicates the relative productivity of the country indicated in the first column to that the USA. The results of the PWT specification are listed in the first column; these estimates correspond to the estimates published in the PWT 9.0. For the second column, labelled "Summed stock", subsoil assets are included as an aggregated production factor, much like capital and labour. We did not include any truncations yet, but this serves to illustrate what happens to the producitivty estimates when subsoil assets are included as a summation over the asset values.

Table 2 – Productivity estimates relative to the USA (USA=1) for 2011

	PWT (1)	Summed stock (2)	Weighted stock (3)
QAT	1.818	1.146	1.05
SAU	1.351	0.939	0.676
USA	1	1	1
NLD	0.912	0.910	0.914
AUS	0.905	0.816	0.762
SWE	0.883	0.889	0.882
CAN	0.807	0.793	0.789
JPN	0.734	0.752	0.745
PRT	0.64	0.652	0.646
URY	0.63	0.644	0.635
ZAF	0.589	0.549	0.527

Note: the first three columns labelled "PWT" show the productivity estimates using PWT 9.0 methodology; "Aggregated Stock" includes the subsoil assets as an aggregate (summed) production factor. "Weighted Stock" shows the results when using our truncation approach introduced above.

The first two columns of Table 4 show that including the subsoil assets in our productivity comparisons is relevant to the estimates for many countries, especially those that are very

[19] The analysis uses 2011 data, as the subsoil asset data is most complete for 2011. Choosing others years, does not impact our results significantly.

dependent on their natural resources. Particularly countries with high shares of subsoil asset income in GDP show declines to their estimated productivity, most notably Qatar (QAT), Saudi Arabia (SAU), and Australia (AUS). On the other hand, countries with very little subsoil assets (income) appear slightly more productive, like Japan (JPN), Portugal (PRT), and Uruguay (URY).

Aggregating and directly comparing different assets, as we do in the second column of Table 2, is only warranted if the marginal products of the aggregated assets are approximately equal. While not directly observable, the marginal products can be approximated by the user costs of capital for each subsoil asset. The user costs for non-renewable capital (like subsoil assets) can be calculated using the Diewert & Fox (2016) approach, which was outlined before. Table 3 shows the user costs for each of the fifteen assets, averaged across countries. The numbers clearly indicate substantial differences between the various assets.

Table 3 – User costs for each type of subsoil asset, averaged across countries for 2011

Asset	User Cost	Asset	User Cost
Copper	0.81	Thermal coal	0.60
Iron ore	0.76	Metallurgical coal	0.51
Gold	0.74	Gas	0.35
Lead	0.69	Bauxite	0.35
Oil	0.67	Tin	0.33
Phosphate	0.65	Silver	0.31
Brown coal	0.61	Zinc	0.30
Nickel	0.61		

Note: the table shows the user cost, or unit rent, as a share of the price of each asset.

To avoid biasing our estimates by grouping assets with different productivities, we disaggregate the stock of subsoil assets into its fifteen constituent parts, and include each separately. With this, the missing factor problem emerges, because not all countries have endowments of all subsoil assets. The third column (labelled "Weighted stock") of Table 2 shows the adjusted estimates using disaggregated stocks of subsoil assets, and our truncation approach. These productivity estimates are based only on the set of common factors between each country, and the USA.

Most of the productivity estimates in the column 3 are smaller than the estimates in column 2, with (again) the most notable differences for Saudi Arabia and Qatar. The difference between these estimates is due to the difference in the composition of the subsoil asset stocks of each country. Specifically, the types of subsoil assets that we now include into the comparison; i.e. only the assets each country pair has in common. The numbers in column 3 are smaller because the stocks of most countries in table 2 consist of rather expensive (in marginal cost terms) resources, like copper, iron ore, gold, and oil. Taking as an example Saudi Arabia, its stock of subsoil assets is dominated by oil stocks, with a high user cost, while that of the USA is much more balanced across different assets. This misalignment causes the productivity of Saudi Arabia to remain upwardly biased in column 2, even though it has already declined relative to PWT estimates. This remaining bias disappears in the third column, because we only compare Saudi Arabia's stock of oil with the USA's stock of oil, rather than the summation of all its subsoil assets.

The change in the estimates between columns 2 and 3 is much less negative for Qatar, and the change relative to that in columns 1 and 2 is much smaller. For the Netherlands (NLD), the change is even positive. This is accounted for by the same effect as the downward movements of the other estimates. The difference is that gas is highly important in both the asset stocks of the Netherlands and Qatar. As per Table 3, gas has one of the lowest user costs. Therefore, for the same reason that Saudi Arabia's estimate decreases, the estimates of Qatar and the Netherlands change much less.

Table 4 shows productivity estimates relative to the USA, which we estimate indirectly, through a third country. These intermediate base countries are Australia (AUS) and Canada (CAN), which we introduced in the data section above, and motivated as being suitable base-countries. We compute the estimates as the relative productivity of any country with an intermediate base country, divided by the relative productivity of the USA compared to the same intermediate country. For each of these two stages, we still need our truncation approach. The differences between the direct (column 1) and the two indirect (columns 2 and 3) comparisons can be quite large. Specifically the comparison with Saudi Arabia and Qatar, again, stand out, as the largest differences between the direct and two indirect productivity estimates.

Table 4 –Productivity Estimates (Weighed stock & USA=1) indirect, linked through Canada (CAN) & Australia (AUS) for 2011

Country	Direct USA	USA via CAN	USA via AUS
QAT	1.050	1.310	0.880
USA	1	1	1
NLD	0.914	0.928	0.778
SWE	0.882	0.88	0.791
CAN	0.789	0.789	0.751
AUS	0.762	0.802	0.763
JPN	0.745	0.777	0.662
SAU	0.676	0.959	0.586
PRT	0.646	0.649	0.683
URY	0.635	0.624	0.548
ZAF	0.527	0.529	0.537

Note: the indirect estimates via AUS (or CAN) are the productivity estimate comparing the comparison country with AUS (CAN), divided by the estimate of the USA relative to AUS (CAN). All estimates in this table correspond to the "Weighted Stock" estimates in Table 2.

We attribute these differences to the structure of the economies in the countries we compare. To illustrate this, we take the example of Qatar, and compare its direct productivity estimate with the USA, to the indirect estimates we show in Table 4. For the direct estimate, we truncate the USA in terms of all its subsoil assets, except for gas and oil, the endowments of Qatar, and estimate their relative productivities. We present this estimate in the final column of Table 2, and for convenience, again in the first column of Table 4.

The differences between direct and indirect estimates occur if specific factors account for very different shares of GDP, in the countries compared. Since the subsoil assets are relatively more important in Australia than in the USA, Canada, and most other countries, estimates linked through Canada are generally closer to the direct estimates. Similarly, the differences are larger for Qatar and Saudi Arabia whose subsoil assets account for sizable shares of GDP (see Table 1). The different sizes of the income shares indicate that the economies are structured differently.[20] The economy of Qatar relies very heavily on subsoil assets and capital as production factors, while many other countries rely much more on their labour force. Such differences in the structure of production make productivity comparisons more difficult, as is revealed through indirect comparisons. Mechanically, the very different factor shares of economies being compared make the weights assigned to each factor of production vary strongly between the first and second stages of the indirect productivity comparison. These changing weights lead the indirect estimates to diverge from the direct productivity estimates.[21]

This result illustrates the difficulty of comparing productivities across very different countries. One way to reconcile this difficulty is by selecting an intermediate base country that is 'in between' the countries being compared, in terms of their factor shares. As an

[20] For South Africa (ZAF), even though its Subsoil assets account for a relatively large share of GDP, the structure of its economy corresponds relatively well to those of Australia, Canada, and the USA. This leads to much smaller differences between direct and indirect comparisons than for example Saudi-Arabia.

[21] When we estimate hypothetical productivity differences for a given pair of countries only using the factor shares of the (intermediate) base country, rather than their own shares, the difference between the direct and indirect estimations disappears.

(imperfect) example, the difference between the direct and indirect estimates of Qatar and Saudi Arabia, are much smaller when we use Australia as the intermediate base country. This is because in Australia, like Qatar and Saudi Arabia, the factor share of labour is lower, capital is higher, and subsoil assets is higher, compared to the USA. Of course, the values of, say, Qatar are more extreme, but those of Australia differ from the USA in the same direction to a greater extent than those of Canada do. In this sense, using Australia is a way to 'bridge' the structural difference between Qatar and Saudi Arabia, and the USA. This is reflected in the indirect estimates being closer to the direct estimates when Australia is used as the intermediate base country. While beyond the scope of this paper, this finding suggests that the use of spanning trees to make international comparisons can improve these comparisons further.[22]

Despite these differences, the indirect comparisons allow us to estimate relative productivities based on the full sets of endowments, even if these sets are not the same. Turning back to the Qatar example, our indirect comparisons allows us to estimate the productivity of Qatar, *based on all its production factors*, relative to the USA, similarly, *based on all its production factors*. This emphasises the difficulty of accurately, and consistently comparing productivities of countries that have organised their production in different ways. By and large, the indirect estimates imply relatively higher productivity levels in the USA when based on the full set of assets, compared to most truncated versions used in the direct estimates.

[22] A spanning trees makes international comparisons by chaining through bilateral comparisons, using only the comparisons with the smallest 'distance' (however defined). See Hill (1999) for more on this method.

Discussion and conclusion

In this paper, we have argued that cross-country productivity comparisons using factors that are missing in some countries, but present in others, lead to biased productivity estimates. To avoid this bias, we propose an industry approach to truncating economies in terms of certain production factors and their associated industries, removing them from the economy. This allows us to estimate productivity differences without the 'missing factor bias'.

Using subsoil assets as an illustration, we have demonstrated our approach to the missing factor endowment problem. We use an industry approach to truncate countries making them consistent for cross-country comparisons across a common set of industries and factors. This way we circumvent the missing endowments problem, and reduced the bias that missing endowments introduced into the results. Furthermore, by linking the estimates through intermediate countries with wide sets of endowments, we estimated productivity comparisons between countries when their endowment sets do not overlap. Finally, our results indicate that including subsoil assets as factors of production can be relevant for productivity estimates across countries, particularly in countries in which these assets constitute an important share of GDP.

Future work could expand our method to cover a broader set of production factors like specific types of labour and capital factors of production. Specifically, as implied by Mutreja (2014), disaggregating the capital stock could yield the finding that marginal productivities of various capital types differ, much like with subsoil assets. Such a finding could imply the possibility of another missing factors problem, this time within the stock of capital production factors.

Furthermore, the possibility of using our truncation approach in combination with the spanning tree method of Hill (1999) offers possibilities for future research to improve further the accurateness of productivity comparisons between countries. Specifically in the situation where the countries being compared have economies that are structured differently.

Additionally, our set of subsoil assets could be expanded into a broader set of natural capital factors, including various types of land, forestry, and environmental factors. While contributing little to solving the missing factors problem, adding such factors could yield interesting insights about the importance of natural capital factors for cross-country productivity differences.

Bibliography

Acemoglu, D., S. Naidu, P. Restrepo and J.A. Robinson, "Democracy does cause growth," Journal of Political Economy, forthcoming, 2016.

Balk, B. M. "On curing the CPI's substitution and new goods bias," CBS, 1999.

Bloom, N., R. Sadun and J. van Reenen, "Management as a technology?," mimeo Stanford University, 2016.

Brandt, N., Schreyer, P., and V. Zipperer, "Productivity measurement with natural capital," Review of Income and Wealth, 63(1): S7-S21, 2017.

Caselli, F., "Accounting for cross-country income differences," In: Handbook of economic growth 1, 679-741, 2005.

Caselli, F., and J. Feyrer, "The marginal product of capital," The quarterly journal of economics, 122(2), 535-568, 2007.

Chen, W., "Cross-Country Income Differences Revisited: Accounting for the Role of Intangible Capital," Review of Income and Wealth, 2017.

Diewert, W. E., “Aggregation problems in the measurement of capital,” In: The measurement of capital, 433-538, University of Chicago Press, 1980.

Diewert, W. E. and C.J. Morrison, "Adjusting Output and Productivity Indexes for Changes in the Terms of Trade," Economic Journal, Royal Economic Society, vol. 96(383), 659-679, September1986.

Diewert, W. E. and K. J. Fox, “The User Cost of Nonrenewable Resources and Green Accounting,” UBC discussion paper, 16-01, 2016.

Feenstra, R. C., “New product varieties and the measurement of international prices,” The American Economic Review, 84(1), 157-177, 1994.

Feenstra, R. C., R. Inklaar, and M. P. Timmer, “The Next Generation of the Penn World Table,” American Economic Review, 105(10): 3150–3182, 2015.

Hill, R. J., “Comparing price levels across countries using minimum-spanning trees,” Review of Economics and Statistics, 81(1), 135-142, 1999.

Hsieh, C.-T. and P. J. Klenow, “Development accounting,” American Economic Journal: Macroeconomics, 2(1), 207-223, 2010.

Inklaar, R. and W.E. Diewert, “Measuring industry productivity and cross-country convergence,” Journal of Econometrics, 191(2): 426–433, 2016.

Inklaar, R. and M. Papakonstantinou, “Vintage effects in human capital: Europe versus the United States,” Review of Income and Wealth, forthcoming, 2018.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman, “Life cycle wage growth across countries,” Journal of Political Economy, 126(2), 797-849, 2018.

Lange, G.-M., Q. Wodon and K. Carey, “The Changing Wealth of Nations 2018,” World Bank: Washington DC, 2018.

Mutreja, P., “Equipment and structures capital: Accounting for income differences,” Economic Inquiry, 52(2), 713-731, 2014.

Redding, S. J., and D.E. Weinstein, “Measuring Aggregate Price Indexes with Demand Shocks: Theory and Evidence for CES Preferences”, 2018.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. de Vries, "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production," Review of International Economics, 23, 575–605, 2015.

Appendix A. Truncating GDP and factor inputs for missing subsoil assets in comparison countries

Subsoil assets are ideally suited for the method presented in this paper, as each asset is generally only used in a single industry, the corresponding mining industry. This means that obtaining industry accounts for our three chosen base countries is required.[23] Specifically, data on value added, labour and capital for each industry, specific to each type of subsoil asset, will enable us to truncate a base country in terms of one or more subsoil assets. Ideally, the industry data covers all the industries associated with each subsoil asset in the World Bank dataset. Such full coverage is not likely to be featured in the data; however, it can be approximated by using more aggregated mining industry data; splitting it according to World Bank output shares.

The industry value added and labour data is obtained from mining censuses, national accounts data, or a combination of both. We obtain the data from each country's national statistical office's website[24]. The capital stock data for each mining industry is estimated from the World Input Output Database (WIOD) database (Timmer et al., 2015).

[23] Note that industry data is only required for the base countries, not for the other (comparison) countries. This is because, with only industry accounts on the base countries, we can create a truncation of each base country to fit each of the comparison countries. This is because for almost all bilateral comparisons, the endowments of the comparison country is a subset of those of the base country (exception: Portugal for USA and Canada).

[24] For Australia the data are obtained from Australian Bureau of Statistics (http://www.abs.gov.au); for Canada, from StatCan (www.statcan.gc.ca); for the USA, the data are from the U.S Census Bureau (https://www.census.gov).

The obtained industry data is used to estimate the value added, capital stocks, and labour stocks of industries specific to each subsoil asset. Firstly, for each industry, the total value added is computed by taking the total value of shipments and services rendered in addition to total capital expenditures, from which the total cost of intermediates is subtracted. This figure is then divided by the total output of the industry to arrive at the value added/output shares for each industry. The detail of the industry classifications varies by country, but in general, the industries correspond to one (or multiple) relevant subsoil assets featured in the World Bank dataset. In the case of multiple industries (and therefore inputs) being combined under a single classification, the overall industry value added/output shares are used for all included subsoil assets. Table A1 displays the value added/output shares for each of the industries, for the three base countries, where missing values correspond to those in Table 1. Note that identical shares of different assets within a country indicate shares taken from more aggregated industry listings.

Table A1 – Value added / Output shares

Asset	AUS	CAN	USA
Oil	0.797	0.662	0.754
Gas	0.797	0.662	0.754
Iron ore	0.786	0.713	0.426
Zinc	0.473	0.766	0.722
Lead	0.473	0.766	0.722
Silver	0.473	0.638	0.702
Copper	0.467	0.766	0.615
Gold	0.449	0.638	0.567
Metallurgical coal	0.441	0.734	0.438
Thermal coal	0.441	0.734	0.438
Brown coal	0.441	0.734	0.562
Phosphate	0.339	0.609	0.59
Nickel	0.307	0.766	
Bauxite	0.307		0.666
Tin	0.307		

The industry labour data shares the same level of aggregation as the value added data (they are from the same source); however, the capital stocks (obtained from WIOD) are estimates for the mining industry as a whole. We therefore assign capital, and where they are aggregated, labour stocks, to each industry according to each subsoil asset's share in World Bank output figures.

The shares listed in table A1 are multiplied with the output figures from the World Bank data, to obtain the value added for each subsoil asset. By subtracting these value added figures from a country's total GDP, we arrive at an estimate of that country's total income without each of the given industries. In addition to this operation however, we need to consider the other factors employed by each industry. The capital and labour that these industries use are also responsible for a part of value that we have previously removed from GDP. It is therefore important that we reduce the country's endowments of capital and labour by the amounts used in the removed industries. Each different truncation, featuring a different set of industries that are removed, will yield different re-estimations of GDP, capital and labour. Table A2 shows for the base countries, the remaining percentage of GDP, capital, and labour after the truncation specific to each bilateral comparison. Since the mining industries are relatively small industries in each of the base countries, and they rely moslty on machinery, the remaining shares of labour hardy drop below 99%. The story is different for GDP and capital. Especially in Australia, GDP and the capital stock are significantly reduced for some comparisons.

Table A2 – The share of truncated GDP and factor inputs

	Australia			Canada			United States		
Country	GDP	Empl.	Capital	GDP	Empl.	Capital	GDP	Empl.	Capital
CAN	100	100	99.7				100	100	100
JPN	93.5	99.1	89.9	98.9	99.8	97.1	99.6	99.9	99.4
NLD	93.1	98.9	88.2	98.7	99.8	96.5	99.5	99.9	99.2
PRT	96.2	99.1	90	94.7	99.5	92.9	98.2	99.8	98.1
QAT	93.1	98.9	88.2	98.7	99.8	96.5	99.5	99.9	99.2
SAU	93.9	99.3	92.4	99.2	99.9	97.9	99.6	100	99.5
SWE	96.6	99.3	91.8	94.8	99.6	93.5	98.2	99.8	98.2
URY	91.9	98.9	88.2	94.3	99.5	92.1	98.1	99.8	98.1
USA	99.9	99.9	99.4	99.8	100	99.5			
ZAF	99.8	99.9	98.9	99.9	100	99.6	99.8	100	99.6